

```
1  GCCCTTGGCA GCAGCCCTGT TACCGCTTAG ATGGCGCGCA GGACAGAGCC
51  CCCCACGGG  GGCTGGGGAC GGGTGGTGGT GCTCTCAGCG TTCTTCCAGT
101 CGGCGCTTGT GTTTGGGGTG CTCCGCTCCT TTGGGGTCTT CTTCGTGGAG
151 TTTGTGGCGG CGTTTGAGGA GCAGGCAGCG CGCGTCTCCT GGATCGCCTC
201 CATAGGAATC GCGGTGCAGC AGTTTGGGAG CCCGGTAGGC AGTGCCCTGA
251 GCACGAAGTT CGGGCCCAGG CCCGTGGTGA TGA CTGGAGG CATCTTGGCT
301 GCGCTGGGGA TGCTGCTCGC CTCTTTTGCT ACTTCCTTGA CCCACCTATA
351 CCTGAGTATT GGGTTGCTGT CAGGCTCTGG CTGGGCTTTG ACCTTCGCTC
401 CGACCCTGGC CTGCCTGTCC TGTTATTTCT CTCGCCGACG ATCCCTGGCC
451 ACCGGGCTGG CACTGACAGG CGTGGGCCTC TCCTCCTTCA CATTTGCCCC
501 CTTTTTCCAG TGGCTGCTCA GCCACTACGC CTGGAGGGGG TCCCTGCTGC
551 TGGTGTCTGC TCTCTCCCTC CACCTAGTGG CCTGTGGTGC TCTCCTCCGC
601 CCACCCTCCC TGGCTGAGGA CCCTGCTGTG GGTGGTCCCA GGGCCCAACT
651 CACCTCTCTC CTCCATCATG GCCCCTTCCT CCGTTACACT GTTGCCCTCA
701 CCCTGATCAA CACTGGCTAC TTCATTCCCT ACCTCCACCT GGTGGCCCAT
751 CTCCAGGACC TGGATTGGGA CCCACTACCT GCCGCCTTCC TACTCTCAGT
801 TGTTGCTATT TCTGACCTCG TGGGGCGTGT GGTCTCCGGA TGGCTGGGAG
851 ATGCAGTCCC AGGGCCTGTG ACACGACTCC TGATGCTCTG GACCACCTTG
901 ACTGGGGTGT CACTAGCCCT GTTCCCTGTA GCTCAGGCTC CCACAGCCCT
951 GGTGGCTCTG GCTGTGGCCT ACGGCTTCAC ATCAGGGGCT CTGGCCCCAC
1001 TGGCCTTCTC TGTGCTGCCT GA ACTAATAG GGACTAGAAG GATTTACTGT
1051 GGCCTGGGAC TGTTGCAGAT GATAGAGAGC ATCGGGGGGC TGCTGGGGCC
1101 TCCTCTCTCA GGCTACCTCC GGGATGTGTC AGGCAACTAC ACGGCTTCTT
1151 TTGTGGTGGC TGGGGCCTTC CTTCTTTCAG GGAGTGGCAT TCTCCTCACC
1201 CTGCCCCACT TCTTCTGCTT CTCAACTACT ACCTCCGGGC CTCAGGACCT
1251 TGTAACAGAA GCACTAGATA CTAAAGTTCC CCTACCCAAG GAGGGGCTGG
1301 AAGGAGGACT GAACTCCACA GAGTCAGGCC CAGAAAGCCA AAGCTTGACA
1351 GCTCCAGGTC TTCTCTTGCC ACGTCTTGGT CTCCACAGAA CCACAGTGCC
1401 TTAAGATTCT TGATCTGCCT CCCCTAGAG CAGGCCTGGG GCTCCTGCAA
1451 TGTGTGTGCC AACCCTTT (SEQ ID NO:1)
```

FEATURES:

5'UTR: 1-30
Start Codon: 31
Stop Codon: 1402
3'UTR: 1405

FIGURE 1A

Docket No.: CL001013CIP-CON
Serial No.: TO BE ASSIGNED
Inventors: KETCHUM, Karen A. et al.
Title: ISOLATED HUMAN TRANSPORTER...

HOMOLOGOUS PROTEINS:

Top 10 BLAST Hits:

	Score	E
CRA 103000001515981 /altid=gi 7670446 /def=dbj BAA95074.1 (AB0...	250	3e-65
CRA 150000165029756 /altid=gi 13431667 /def=sp O70461 MOT3_RAT ...	244	1e-63
CRA 89000000192725 /altid=gi 10048452 /def=ref NP_065262.1 sol...	238	8e-62
CRA 18000005042369 /altid=gi 2497855 /def=sp Q63344 MOT2_RAT MO...	238	1e-61
CRA 18000005039313 /altid=gi 1432167 /def=gb AAB04023.1 (U6231...	238	1e-61
CRA 18000005141743 /altid=gi 6755536 /def=ref NP_035521.1 solu...	234	2e-60
CRA 335001098681302 /altid=gi 11418102 /def=ref XP_009979.1 mo...	234	2e-60
CRA 1000682335761 /altid=gi 7019529 /def=ref NP_037488.1 monoc...	233	5e-60
CRA 18000005141744 /altid=gi 4759120 /def=ref NP_004722.1 solu...	232	6e-60
CRA 108000024650708 /altid=gi 12737028 /def=ref XP_012127.1 so...	232	6e-60

BLAST dbEST hits:

	Score	E
gi 8423571 /dataset=dbest /taxon=960...	733	0.0

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

From BLAST dbEST hits:

gi|8423571 breast

From tissue screening panels:

Spleen

Breast (adult)

FIGURE 1B

1 MARRTEPPDG GWGRVVVLSA FFQSALVFGV LRSFGVFFVE FVAAFEEQAA
51 RVSWIASIGI AVQQFGSPVG SALSTKFGPR PVVMTGGILA ALGMLLASFA
101 TSLTHLYLSI GLLSGSGWAL TFAPTLACLS CYFSRRRSLA TGLALTGVGL
151 SSFTFAPFFQ WLLSHYAWRG SLLLVSALS LHLVACGALLR PPSLAEDPAV
201 GGPRAQLTSL LHHGPFLRYT VALTLINTGY FIPYLHLVAH LQDLWDPLP
251 AAFLLSVVAI SDLVGRVVS G WLGDAVPGPV TRLLMLWTTL TGVSLALFPV
301 AQAPTALVAL AVAYGFTSGA LAPLAFSVLP ELIGTRRIYC GLGLLQMIES
351 IGLLGPPLS GYL RDVSGNY TASFV VAGAF LLSGSGILLT LPHFFCFSTT
401 TSGPQDLVTE ALDTKVPLPK EGLEGGLNST ESGPESQSLT APGLLLPRLG
451 LHRTTVP (SEQ ID NO:2)

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 2

1 369-372 NYTA
2 428-431 NSTE

[2] PDOC00004 PS00004 CAMP_PHOSPHO_SITE
cAMP- and cGMP-dependent protein kinase phosphorylation site

135-138 RRRS

[3] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 3

1 74-76 STK
2 134-136 SRR
3 335-337 TRR

[4] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 2

1 193-196 SLAE
2 432-435 SGPE

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 18

1 29-34 GVLRSF
2 66-71 GSPVGS
3 70-75 GSALST
4 86-91 GGILAA
5 87-92 GILAAL
6 93-98 GMLLAS
7 111-116 GLLSGS
8 115-120 GSGWAL
9 142-147 GLALTG
10 147-152 GVGLSS
11 201-206 GGPRAQ
12 292-297 GVSLAL
13 368-373 GNYTAS
14 386-391 GILLTL
15 422-427 GLEGGL

FIGURE 2A

16 425-430 GGLNST
17 426-431 GLNSTE
18 450-455 GLHRTT

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	13	33	1.302	Certain
2	52	72	1.039	Certain
3	81	101	2.101	Certain
4	114	134	1.703	Certain
5	139	159	1.850	Certain
6	170	190	1.572	Certain
7	219	239	1.192	Certain
8	245	265	1.019	Certain
9	283	303	1.832	Certain
10	306	326	1.709	Certain
11	338	358	0.976	Putative
12	372	392	1.982	Certain

FIGURE 2B

Docket No.: CL001013CIP-CON
Serial No.: TO BE ASSIGNED
Inventors: KETCHUM, Karen A. et al.
Title: ISOLATED HUMAN TRANSPORTER...

BLAST Alignment Top Hit:

>CRA|150000165029756 /altid=gi|13431667 /def=sp|O70461|MOT3_RAT
MONOCARBOXYLATE TRANSPORTER 3 (MCT 3) /org=MCT 3
/dataset=nraa /length=492
Length = 492

Score = 244 bits (617), Expect = 1e-63
Identities = 168/470 (35%), Positives = 239/470 (50%), Gaps = 36/470 (7%)

Query: 3 RRTEPPDGGWGRVVVLSAFFQSALVFGVLRSGVFFVEFVAAFEQAAARVSWIASIGIAV 62
R PPDGGWG VV+ + F + +G ++ VFF E F + +W++SI +A+
Sbjct: 8 RGAGPPDGGWGWVVLGACFVITGFAYGFPAKAVSVFFRELKRDFGAGYSDTAWVSSIMLAM 67

Query: 63 QQFGSPVGSALSTKFGPRPVMTGGILAAALGMLLASFATSLTHLYLSIGLLSGSGWALTF 122
P+ S L T+FG RPV++ GG+LA+ GM+LASFA+ L LYL+ G+L+G G AL F
Sbjct: 68 LYGTGPLSSILVTRFGCRPVMLAGLLASAGMILASFASRLLELYLTAGVLTGLGLALNF 127

Query: 123 APTLACLSCYFSRRRSLATGLALTGVGLSSFTFAPFFQWLLSHYAWRGSLLLVSALS LHL 182
P+L L YF RRR LA GLA G + T +P Q L + WRG LL L LH
Sbjct: 128 QPSLIMLGlyfERRRPLANGLAAAGSPVFLSTLSPLGQLLGERFGWRGGFLLFGGLLLHC 187

Query: 183 VACGALLRPPSLAE---DPAVGGPRAQLTSLH-----HGPFLRYTVALTLINTGYFIPY 234
ACGA++RPP + DPA G RA+ LL F+ Y V L+ G F+P
Sbjct: 188 CACGAVMRPPPGPQRPDPAPPGRARHRQLLDLAVCTDRTFMVYMTKFLMALGLFVPA 247

Query: 235 LHLVAHLQDLWDPLPAAFLLSVVAISDLVGRVVGWLG--DAVPGPVTRLLMLWTTLTG 292
+ LV + +D AAFLLS+V D+V R G L + V L L G
Sbjct: 248 ILLVNYAKDAGVPDAEAAFLLSIVGFVDIVARPACGALAGLGRLRPHVPYLFSLALLANG 307

Query: 293 VSLALFPVAQAPTALVALAVAYGFTSGALAPLAFSVLPELIGTRRIYCGLGLLQMIESIG 352
++ + A++ LVA +A+G + G + L F VL +G R LGL+ ++E++
Sbjct: 308 LTDLISARARSYGTLVAFCLAFGLSYGMVGAHQFEVLMATVGAPRFPSALGLVLLVEAVA 367

Query: 353 GLLGPPLSGYLRDVSNGNYTASFVVAGAFLLSGSGILLTLPHFFCFSTT----- 400
L+GPP +G L D NY F +AG+ ++ +G+ + + C +
Sbjct: 368 VLIGPPSAGRLVDALKNYEIIIFYLAGS-EVALAGVFMAVTTYCCLRCSKNISSGRSAEGG 426

Query: 401 TSGPQDLVTEALDTKVPLPKEGLEGLNSTESGPESQSLTAPGLLLPRLG 450
S P+D+ EA P+P STE E SL A +L PR G
Sbjct: 427 ASDPEDV--EAERDSEPMPA-----STE---EPGSLEALEVLSPRAG 463 (SEQ ID
NO:4)

>CRA|89000000192725 /altid=gi|10048452 /def=ref|NP_065262.1| solute
carrier family 16 (monocarboxylic acid transporters),
member 8; proton-coupled monocarboxylate transporter 3
gene; proton-coupled monocarboxylate transporter 3 [Mus
musculus] /org=Mus musculus /taxon=10090 /dataset=nraa
/length=492
Length = 492

Score = 238 bits (602), Expect = 8e-62
Identities = 165/470 (35%), Positives = 236/470 (50%), Gaps = 36/470 (7%)

Query: 3 RRTEPPDGGWGRVVVLSAFFQSALVFGVLRSGVFFVEFVAAFEQAAARVSWIASIGIAV 62
R PPDGGWG VV+ + F + +G ++ VFF E F + +W++SI +A+
Sbjct: 8 RGAGPPDGGWGWVVLGACFVVTGFAYGFPAKAVSVFFRELKRDFGAGYSDTAWVSSIMLAM 67

Query: 63 QQFGSPVGSALSTKFGPRPVMTGGILAAALGMLLASFATSLTHLYLSIGLLSGSGWALTF 122
P+ S L T+FG RPV++ GG+LA+ GM+LASFA+ L LYL+ G+L+G G AL F

FIGURE 2C

Docket No.: CL001013CIP-CON
Serial No.: TO BE ASSIGNED
Inventors: KETCHUM, Karen A. et al.
Title: ISOLATED HUMAN TRANSPORTER...

Sbjct: 68 LYGTGPLSSILVTRFGCRPVMLAGGLLASAGMILASFASRLVELYLTAGVLTGLGLALNF 127

FIGURE 2D

Docket No.: CL001013CIP-CON
Serial No.: TO BE ASSIGNED
Inventors: KETCHUM, Karen A. et al.
Title: ISOLATED HUMAN TRANSPORTER...

Query: 123 APTLACLSCYFSRRRSLATGLALTGVGLSSFTFAPFFQWLLSHYAWRGSLLLVSALS LHL 182
P+L L YF RRR LA GLA G + +P Q L + WRG LL L LH
Sbjct: 128 QPSLIMLGLYFERRRPLANGLAAAGSPVFLSMLSPLGQLLGERFGWRGGFLLFGGLLLHC 187

Query: 183 VACGALLRP---PSLAEDPAVGGPRAQLTSLH-----HGPFLRYTVALTLINTGYFIPY 234
ACGA++RP P DP+ G A+ LL F+ Y V L+ G F+P
Sbjct: 188 CACGAVMRPPPGPPRRDPSPHGGPARRRRLLDVAVCTDRAFFVYVVTKFLMALGLFVPA 247

Query: 235 LHLVAHLQDLWDPLPAAFLLSVVAISDLVGRVVSGLG--DAVPGPVTRLLMLWTTLTG 292
+ LV + +D AAFLLS+V D+V R G L + V L L G
Sbjct: 248 ILLVNYAKDAGVPDAEAAFLLSIVGFVDIVARPACGALAGLRRLRPHVPYLFSLALLANG 307

Query: 293 VSLALFPVAQAPTALVALAVAYGFTSGALAPLAFSVLPELIGTRRIYCGLGLLQMIESIG 352
++ + A++ LVA +A+G + G + L F VL +G R LGL+ ++E++
Sbjct: 308 LTDLISARARSYGTLVAFICIAFGLSYGMVGALQFEVLMATVGAPRFPSALGLVLLVEAVA 367

Query: 353 GLLGPPLSGYL RDVSGNYTASFVVGAFLLSGSGILLTLPHFFCFSTT----- 400
L+GPP +G L D NY F +AG+ ++ +G+ + + + C +
Sbjct: 368 VLIGPPSAGRLVDALKNYEIIIFYLAGS-EVALAGVFMAVTTYCCLRCSKNISSGRSAEGG 426

Query: 401 TSGPQDLVTEALDTKVPLPKEGLEGLNSTESGPESQSLTAPGLLLPRLG 450
S P+D+ EA P+P STE E SL A +L PR G
Sbjct: 427 ASDPEDV--EAERDSEPMPA-----STE---EPGSLEALEVLSPRAG 463 (SEQ ID NO:5)

Hammer search results (Pfam):

Model	Description	Score	E-value	N
PF01587	Monocarboxylate transporter	204.9	1.2e-57	2
PF01925	Domain of unknown function	4.4	4.6	1
PF00348	Polyprenyl synthetases	3.7	6.1	1
PF00083	Sugar (and other) transporter	3.0	3.8	1
PF01306	LacY proton/sugar symporter	2.7	6.6	1
PF01309	Equine arteritis virus small envelope glycop	2.3	5	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF01925	1/1	65	97 ..	165	201 .]	4.4	4.6
PF00083	1/1	12	108 ..	1	113 [. .	3.0	3.8
PF01309	1/1	153	173 ..	1	21 [. .	2.3	5
PF00348	1/1	174	191 ..	1	19 [. .	3.7	6.1
PF01587	1/2	20	192 ..	1	191 [. .	160.8	2.3e-44
PF01587	2/2	219	377 ..	441	611 .]	48.3	1.6e-12
PF01306	1/1	373	393 ..	393	415 .]	2.7	6.6

FIGURE 2E

```
1  CATT TTTTAGT GCATGGATTT TCTAACTGAA CCCCTTGGGC AACGCTTAAT
51 AGTAGGTACT ATTATCCCCA GTTTACAGAT GGGGAAACCA ACTGAGAGAT
101 TCAGCATCTT GATCGAGTTA AGTAATAAAG TCAAGATTGG AACTGGGCCA
151 GGCACGGTGG CTCACGCCTG TAATCCCAGC ACTTTGGGAG GCCAAGGCTG
201 GTGGATCACT TGAGGTCAGG AGTTCGAGAC CAGCGTGGCC AACATGGTGA
251 GACCTCGTCT CTAATAAAAA TACCAAAATT AACTGGGCGT TGTGGTGGGA
301 GCCTGTAATC CCAGAACTC AGGAGACTGA GGCAGGAGAA TCACTTGAAC
351 CCGGGAGGTG GAGGTGTCAG TGAGCCAAGA TCATGCCACT GCACTCCAGC
401 CTGGGCCACA GAGCAAGACT CCGTCTCAAA ATAAATAAAT AAATAAATAA
451 ATAAATAAAA GACTGGAAC GTGATCTGAT TCTAAAGACC CGAGTTCTTA
501 ATCACTATGT AATACAGCCA CAGCAATTTT TGTATCTTTG GCATATTTCC
551 CACCAGCCGA CATT TTTGACT CTTAGAAAGT ATATATGTGT ATTATTGATG
601 ATTACTTTTA TTTCCACAT ATAAAATTAT TTAAGGCTCA ATATGTCTTT
651 TAAGACTGCA CACCTCCCTC CCTGCCTCCA CTTCTTGTTT GCTGCTTTCC
701 CCAGTAATCT GGGAGTGAAC ATTGAGTCCA CGGTTTCAAG GTCAGGGTCC
751 TGGGAAGTAT GGCTTATAAT GAAGGAACAG GAAATCCAAG CCATTGGTGT
801 TATGGAGACT GGGGAAGGACT GGGGAGTGTT TGCTAGGGGC CTGAGGACTA
851 CTTGGGTAA GAGGGGCTGA CTGCTCCAGT GGCCAGGGTC ATAGTTTGT
901 CTCTTTAGTC TACCCACCA TCAGATCAAA AAAGGTGGTT AGGAAGTGGT
951 TGTTACTAGA GGGCAGAGGA AAAGGTCCA GCCCCAGTGA GGAAGAGGTA
1001 GGTGGTGTG GTGGGGCCCT GTGTGAGCTT ACAGCCGCCC TTCCTCTCCT
1051 CAGTTATTTT TGGTCTCTGT GACCTGTAGG TTTCTGTGTA GTGGGAACAG
1101 AAGTGACAGG AACGAGTTCC CACTACAGAA ATGAACGCCA GGAGTCCAAC
1151 TCATTCCCCT TCTCTCTTCC CTTAGCCGTT GAACTTCTCA GGGATCCAGG
1201 CTTCTAGGTC TGCGTGCCTA GGGCTGCGTG TTAGTGGCTT CAGGCGCTGC
1251 GCCAAACACT TCGTTTGAGT CTCATCTCCT AACCCCTCCC CTACCCCAA
1301 CAGGGCCTTG CAATTCCTGG ACCCTCATT AAAGCAAGAG AGTCCTCTCC
1351 TCTCCAGACC CAGTTTACCC ACCACTAACC CTTCCGTGTG GCTCTGGGTG
1401 CTGAAACGGG GATGACTTGG CCCGCTAGGT GAAGAGGAGA CGGAAGCTTC
1451 CTGGCAGTCC CCGCGTCACG TGGGGCCCTA CCTAGTCAGC CTCCTAACGC
1501 CCCTCCTTAC GCATGCGCCC ATTCAGTGT GGTCCCCAAC AATGCCTAAA
1551 TCCCGCCCTG CCCTTCTCGT TCCGCCCCTG CCCGGGAGCC CCGCGTCTCT
1601 ATTGGCGAGC TCCAGGGTGG CCCGGCCCCG ACACCCAGT GATAAAATAG
1651 ATCATCTACA CGGAACTGG CGCGCTCCAG GGGTGGGGCC CAACTCAGT
1701 TCCACCCTCT GGCTCCCAGC CGAACACCGA ACCGGGACCG ATCCGGCCCC
1751 GGCTTGAAC AGCTCAGCTC CGAGCTCGCG GAACCACGCC CCCGGGAGAC
1801 TCTGGCCCCG CCAGCGCGGG CCAGGTCTTC AGTCCTATAT CGCCCTGCCT
1851 TGGGAAAGG TGCAGGGGCC TCTCGCCGCC TCGTCGGGCC CTTCTCTCT
1901 ACCTGCCTCT CCAACCCCTC TCGCCCCGA GCCACCCGGC AGCGGGGGTG
1951 GGTGTGCAGA GGTGCGGCGT CCAGAACCCG GCTCCTGCAG AGGCTCTGGG
2001 TGGCAGCAGC CCTGTTACCG CTTAGATGGC GCGCAGGACA GAGCCCCCG
2051 ACGGGGGCTG GGGATGGGTG GTGGTGTCT CAGCGTTCTT CCAGTCGGCG
2101 CTTGTGTTT GGGTGCTCCG CTCCTTTGGG GTCTTCTTCG TGGAGTTTGT
2151 GCGGCGGTTT GAGGAGCAGG CAGCGCGCGT CTCCTGGATC GCCTCCATAG
2201 GAATCGCGGT GCAGCAGTTT GGGAGTGAGT GCGGCGCCTG GATCTGGCGG
2251 ACTGCGACCC TCGGAAGGGA GAGGGAATGC GCGACTGGG AAGTGGAAGG
2301 GCGAGGGGCG GGAGATGCTG GGGGGGAGAC CCCTGAGATC TTCTCGCAGC
2351 GCCCCTTCCA CTTCTCAGG CCCGGTAGGC AGTGCCCTGA GCACGAAGTT
2401 CGGGCCCAGG CCCGTGGTGA TGACTGGAGG CATCTTGGCT GCGCTGGGGA
2451 TGCTGCTCG CTCTTTTGCT ACTTCCTTGA CCCACCTATA CTTGAGTATT
2501 GGGTTGCTGT CAGGTGAGAG CCTGCACAAG GGCAGGAGAG TCAAATGCTT
2551 AGATCGTTGG ATGTTACCT CTTCTCTGCT CTTCCAAAG GGTTCGGGGA
2601 GAAGCTGAGG GAAAGTTTAG CTAGCACCTG TACCCAGAAG GGAATTCTTA
2651 ATAGGAATGA CTAAAGCGAC AAACATGGTG AGGAATTAGG AAATTCAAGG
2701 ATGATGAAAC CTGGCCAGGC ACGGTGGCTC ACGCCTGTAA TCCCAGCACT
2751 TTGGGAAGCC GAGGCGGGTG GATCACGAGG TCAGGAGTTT GAGACCAGCC
2801 TGGCCAACAT GGTGAAACCC CGTCTCTACA AAAATACAAA AATTAGCCGG
2851 GCCTGGTGGC GCTAATCCCA GTTACTCGGG AGGCTGAGGC AGGAGAATCG
2901 CTTGAACCCG GGAGGCGGAG GTTGCAAGTA GCCAAGATCG CACCACTGCA
2951 CTCCAGCCTG GGCGACAGAG CAAGATTCTG TCTCAAAAAA AAAAAAAA
```

FIGURE 3A


```

3001 AAAAAAAAAA AGATGAAACC AAGTATACAA GCCCAGAAGC CTAGGGCTAA
3051 TGGGACTGGA GTGCAAAAGG AAGAATTACT ATAAAATGGT GCTAGGGGCC
3101 AGGCACGGTG GCTCACGCCT GTAATCCCAG CACTTTGGGA GGCCGAGGCG
3151 GGCGGATCAC GAGGTCAGGA GATCAAGACC ATCCTGGCTA ACACGGTGAA
3201 ATCACGTCTC TACTAAAAAC ACAAAAAATT AGCTGGGCGT GGTGGCAGGT
3251 GACTGTAGTC CCAGCTACTC GGGAGGCTGA GGCAGGAGAA TGGTGTGAAC
3301 CCGGGAAGCA GAGCTTGCA GAGCCGAGA TTGCACCACT GCACTCCAGC
3351 CTGGGCGACA GAGCGAGACT CCGTCTCAAA AAAAAAAGA AAAAAAAGG
3401 TGCTAGGTAC TGTGACTGTG AAATCGATAT CATTATTGGA TTTACAGCTG
3451 GGGAAAAGCT TTAAAGCTTA TACAACTTGG CAAATGAAGG TCACACAGCT
3501 AGAAATGGTA GAGCCCAGGT CTAACCTCAA AGTTCTGTGC TAGTTACCTT
3551 ACAAACTTTG TCTCTAATCT TCCACAATCC CAAAAAGTGT ATTATTACAT
3601 TTTGCAGTTG AGAAGGTTGA GGCTGGGGGT GTTAAGTAAA ACACACAAGG
3651 TTACACAGCT ATGAAGTATC CAAGCCAAGA TTGTATCCCA GGTCTGTGGG
3701 ACTCCGAAGC AAGTGCTACA TTCTGCTGCT GGGCAATGCG GGGATTACTG
3751 TGTGCCCTGA GCTCCCTAAG AGTTCTCAAC ACCACTTCTT CCTTTTGTAC
3801 AGGCTCTGGC TGGGCTTTGA CCTTCGCTCC GACCCTGGCC TGCTGTCTCT
3851 GTTATTTCTC TCGCCGACGA TCCCTGGCCA CCGGGCTGGC ACTGACAGGC
3901 GTGGGCCTCT CCTCCTTCAC ATTTGCCCCC TTTTTCAGT GGCTGCTCAG
3951 CCACTACGCC TGGAGGGGGT CCCTGCTGCT GGTGTCTGCC CTCTCCCTCC
4001 ACCTAGTGGC CTGTGGTGCT CTCTCCGCC CACCCTCCCT GGCTGAGGAC
4051 CCTGCTGTGG GTGGTCCAG GGCCCAACTC ACCTCTCTCC TCCATCATGG
4101 CCCCTTCCTC CGTTACACTG TTGCCCTCAC CCTGATCAAC ACTGGCTACT
4151 TCATTCCCTA CCTCCACCTG GTGGCCCATC TCCAGGACCT GGATTGGGAC
4201 CCACTACCTG CTGCCTTCCT ACTCTCAGTT GTTGCTATTT CTGACCTCGT
4251 GGGGCGTGTG GTCTCCGGAT GGCTGGGAGA TGCAGTCCCA GGGCCTGTGA
4301 CACGACTCCT GATGCTCTGG ACCACCTTGA CTGGGGTGTC ACTAGCCCTG
4351 TTCCCTGTAG CTCAGGCTCC CACAGCCCTG GTGGCTCTGG CTGTGGCCTA
4401 CGGCTTCACA TCAGGGGCTC TGGCCCCACT GGCCTTCTCT GTGCTGCCTG
4451 AACTAATAGG GACTAGAAGG ATTTACTGTG GCCTGGGACT GTTGCAGATG
4501 ATAGAGAGCA TCGGGGGGCT GCTGGGGCCT CCTCTCTCAG GTAAGTGGAA
4551 TGGGGTTCCC AGGGGGTGAG GGCTGCCATG TTGCACAACCT AGGGGAGGGT
4601 ACTATTCTCA TTACAGTGTA TGTGAATATT GCCCTCTGGT GTAGTACAGT
4651 ACACAGCCTG CGTGGCCAAC CATAGCATCC CTGAAATGGG TCCATGGGGC
4701 AAAGAACTTG GGGCTGGGAA AGTCTGAGTG GAAAGACAAA AAGAAGCTAA
4751 GTGGAACCTT TGGCAGGGTG CCTACGGCTT GGGTTTGCAG AGGACCTGGC
4801 AGAACCTGGC CAGACACAGA CGTAGCATTC CAGTGTGCAC CCTTTCCTTT
4851 GGCCTACTGG GCCCCAAACC AGGTATCTGA GGCACCTGGT CAAAGTTCTG
4901 CTGGCTCAGG GTGCCAGAAC TTTCAGACCT TTATCTCCTC TTACCCATTA
4951 ACTGAAGCTT TAGAAAGGCC ACAGTTGGTG GGCCTCTGTA GTCCCAGCTA
5001 CTCAGGAGGC TGAGGCAGGA GAATGGCATG AACC CGGAG GCGGAGCTTG
5051 CAGTGAGCTG AGATCGCGCC ACTGCACCTC AGCCTGGGCG ACAGAGCGAG
5101 ACTCCGTCTC AAAAAAAAAA AAAAAAGAAA GGCCACAGTT GCCAGAAAGA
5151 AAGGCACAAG TATGCCTGAC TCAATCTGGA TCTCCAAATC CCTGCAGGCT
5201 GGTTTGGAGG TCCTTTCTGA AGGCGGGGAG GTGGTTGAAA TTAACTTTGT
5251 AGGCCCTTTT GGGAAACCAG AGTTCTTAAG TTTATCCAAC TATCCATGG
5301 GAGTTCCAAC TCCTCTGAGA TGATAAGTCT TCCCTCCACC CAAAAATGTA
5351 TCTGAGCCCT CAGCCCCAGC AAATAGATCA CTCATGTGTA TTCTTTTCTT
5401 CTCTTGACC TAGGCTACCT CCGGATGTG ACAGGCAACT ACACGGCTTC
5451 TTTTGTGGTG GCTGGGGCCT TCCTTCTTTC AGGGAGTGGC ATTCTCCTCA
5501 CCCTGCCCCA CTTCTTCTGC TTCTCAACTA CTACCTCCGG GCCCCAGGAC
5551 CTTGTAACAG AAGCACTAGA TACTAAAGTT CCCCTACCCA AGGAGGGACT
5601 GGAAGGAGGA CTGAACTCCA CAGAGTCAGG CCCAGAAAGC CAAAGCTTGA
5651 CAGCTCCAGG TCTTCTCTTG CCACGTCTTG GTCTCCACAG AACCACAGTG
5701 CCTTAAGATT CTTGATCTGC CTCCCCCTAG AGCAGGCCTG GGGCTCCTGC
5751 AATGTGTGTG CCAACCTTTT GTATTTTGTG GAGGACTCTT ATTTCTCCGT
5801 TACTCTCCTA ACCTTTTCTT CTTTTTCTT TTTCCCGAGA CGGAGTCTTG
5851 CTCTGTTGCC CAGGCTGGAG TGCACTGATG TGATCTCGGC TCACTGCAAC
5901 CTCCGCTTCC CGGGTTCAAG CGATTCTCCT GCCTCAGCCT CCCAAGTAGC
5951 TGGGATTACA GGCGGGAGCC ACCACACCCG GCTATTTTTT TTTTTTTTTT

```

FIGURE 3B

```

6001 TTTNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNTTTTGG TAGAGACAGG
6051 GTTTCACCAT GTTGGCCAGG ATGGTCTCGA ACTCCTGACC TTGTGATCCA
6101 CCCCCGCCC CTCCCTCGGC CTTCCAAAGT GCTGGGATTA CAGGCGTGAG
6151 CCACCACACC CAGCCTCCCC TAACCTTTTC TAAAGGACCC AGGAGTTTTC
6201 AAGGATCCGG GAGTTCCTGC TTCCTGAGC TGTGAATCAA CTGTGAAAAT
6251 CAAAGGCCAA GAGACTTATC ATGCTTTATA TAACATCTCT AGTGTGTCCT
6301 CCTGAGTTTC TTCTCTGAAG ACACATGTTT GGGAAACAAA ACTGTCCCTT
6351 TGAGATAAAA TCAAATAAGA AAATTGGATA ATAATCACAA CCTCAAAATG
6401 AGCTGGGGCC CATATGCTTG GGTGGCCGA ATGGAGTCAT GCCTGGAAGT
6451 GGAGGAGAGT GTCCAGGAGC TCCGATGACC CAAGGCATTT TAACCCTGGA
6501 ATCTGCTCTC CAGGCTACCA CCACATACCT CCCTCTTCCC CATTATCCCT
6551 GTGGCTTAGA AAAGAA (SEQ ID NO:3)

```

FEATURES:

Start: 2026
Exon: 2026-2224
Intron: 2225-2369
Exon: 2370-2513
Intron: 2514-3802
Exon: 3803-4540
Intron: 4541-5413
Exon: 5414-5703
Stop: 5704

CHROMOSOME MAP POSITION:

Chromosome 17

ALLELIC VARIANTS (SNPs):

DNA				Protein		
Position	Major	Minor	Domain	Position	Major	Minor
423	G	A	Beyond ORF (5')			
2717	A	G	Intron			
3064	C	T	Intron			
4146	C	A	Exon	229	G	G
4440	T	C	Exon	327	S	S
4443	G	T	Exon	328	V	V
5105	T	C	Intron			

Context:

DNA

Position

```

423 TAATAAAGTCAAGATTGGAAGTGGGCCAGGCACGGTGGCTCACGCCTGTAATCCCAGCAC
TTTGGGAGGCCAAGGCTGGTGGATCACTGAGGTCAGGAGTTCGAGACCAGCGTGGCCAA
CATGGTGAGACCTCGTCTCTACTAAAAATACCAAATTAAGTGGGCGTTGTGGTGGGAGC
CTGTAATCCCAGAACTCAGGAGACTGAGGCAGGAGAATCACTTGAACCCGGGAGGTGGA
GGTTGCAGTGAGCCAAGATCATGCCACTGCACTCCAGCCTGGGCCACAGAGCAAGACTCC
[G,A]
TCTCAAAATAAATAAATAAATAAATAAATAAATAAAGACTGGAAGTGTGATCTGATTCT
AAAGACCCGAGTTCTTAATCACTATGTAATACAGCCACAGCAATTTCTGTATCTTTGGCA
TATTCACCCAGCCGACATTTTGACTCTTAGAAAGTATATATGTGTATTATTGATGATT
ACTTTTATTTCCACATATAAAATTATTTAAGGCTCAATATGTCTTTTAAGACTGCACAC
CTCCCTCCCTGCCTCCACTTCTGTTTGCTGCTTTCCCAGTAATCTGGGAGTGAACATT

2717 GTGATGACTGGAGGCATCTTGGCTGCGCTGGGGATGCTGCTCGCCTCTTTTGCTACTTCC
TTGACCCACCTATACCTGAGTATTGGGTGCTGTGAGGTGAGAGCCTGCACAAGGGCAGG
AGAGTCAAATGCTTAGATCGTTGGATGTTACCTCCTTCCTGCTCCTTCAAAGGGTTCG
GGGAGAAGCTGAGGGAAAGTTTAGCTAGCACCTGTACCCAGAAGGGAATTCTTAATAGGA

```

FIGURE 3C

Docket No.: CL001013CIP-CON
Serial No.: TO BE ASSIGNED
Inv ntors: KETCHUM, Karen A. et al.
Title: ISOLATED HUMAN TRANSPORTER...

ATGACTAAAGCGACAAACATGGTGAGGAATTAGGAAATTCAAGGATGATGAAACCTGGCC
[A, G]
GGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAAGCCGAGGCGGGTGGATCACG
AGGTCAGGAGTTTGAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACAAAATAC
AAAAATTAGCCGGGCCTGGTGGCGCTAATCCCAGTTACTCGGGAGGCTGAGGCAGGAGAA
TCGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATCGCACCACTGCACTCCAGC
CTGGGCGACAGAGCAAGATTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAGATGAA

FIGURE 3D

3064 GCGGGTGGATCACGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACATGGTGAAACCCCGT
CTCTACAAAAATACAAAAATTAGCCGGGCTGGTGGCGCTAATCCCAGTTACTCGGGAGG
CTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATCGCAC
CACTGCACTCCAGCCTGGGCGACAGAGCAAGATTCTGTCTCAAAAAAAAAAAAAAAAAA
AAAAAAAAAGATGAAACCAAGTATACAAGCCCAGAAGCCTAGGGCTAATGGGACTGGAGTG
[C, T]
AAAAGGAAGAATTACTATAAAATGGTGCTAGGGGCCAGGCACGGTGGCTCACGCCTGTAA
TCCCAGCACTTTGGGAGGCCGAGGCGGGCGGATCACGAGGTCAGGAGATCAAGACCATCC
TGGCTAACACGGTGAAATCACGTCTCTACTAAAAACACAAAAAATTAGCTGGGCGTGGTG
GCAGGTGACTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGTGTGAACCCGG
GAAGCAGAGCTTGCACTGAGCCGAGATTGCACCACTGCACTCCAGCCTGGGCGACAGAGC

4146 GTCCTGTTATTTCTCTCGCCGACGATCCCTGGCCACCGGGCTGGCACTGACAGGCGTGGG
CCTCTCCTCCTTACATTTGCCCCCTTTTCCAGTGGCTGCTCAGCCACTACGCCTGGAG
GGGGTCCCTGCTGCTGGTGTCTGCCCTCTCCCTCCACCTAGTGGCCTGTGGTGTCTCCT
CCGCCACCCCTCCCTGGCTGAGGACCCTGCTGTGGGTGGTCCCAGGGCCCACTCACCTC
TCTCCTCCATCATGGCCCCCTTCTCCGTTACACTGTTGCCCTCACCTGATCAACACTGG
[C, A]
TACTTCATTCCCTACCTCCACCTGGTGGCCCATCTCCAGGACCTGGATTGGGACCCACTA
CCTGCTGCCTTCCTACTCTCAGTTGTTGCTATTTCTGACCTCGTGGGGCGTGTGGTCTCC
GGATGGCTGGGAGATGCAGTCCCAGGGCCTGTGACACGACTCCTGATGCTCTGGACCACC
TTGACTGGGGTGTCACTAGCCCTGTTCCCTGTAGCTCAGGCTCCCACAGCCCTGGTGGCT
CTGGCTGTGGCCTACGGCTTCACATCAGGGGCTCTGGCCCCACTGGCCTTCTCTGTGCTG

4440 CACTGGCTACTTCATTCCCTACCTCCACCTGGTGGCCCATCTCCAGGACCTGGATTGGGA
CCCACTACCTGCTGCCTTCCTACTCTCAGTTGTTGCTATTTCTGACCTCGTGGGGCGTGT
GGTCTCCGGATGGCTGGGAGATGCAGTCCCAGGGCCTGTGACACGACTCCTGATGCTCTG
GACCACCTTGACTGGGGTGTCACTAGCCCTGTTCCCTGTAGCTCAGGCTCCCACAGCCCT
GGTGGCTCTGGCTGTGGCCTACGGCTTCACATCAGGGGCTCTGGCCCCACTGGCCTTCTC
[T, C]
GTGCTGCCTGAACTAATAGGGACTAGAAGGATTTACTGTGGCCTGGGACTGTTGCAGATG
ATAGAGAGCATCGGGGGGCTGCTGGGGCCTCCTCTCTCAGGTAAGTGGAATGGGGTTCCC
AGGGGGTGAGGGCTGCCATGTTGCACAAGTAGGGGAGGGTACTATTCTCATTACAGTGTA
TGTGAATATTGCCCTCTGGTGTAGTACAGTACACAGCCTGCGTGGCCAACCATAGCATCC
CTGAAATGGGTCCATGGGGCAAAGAACTTGGGGCTGGGAAAGTCTGAGTGGAAGACAAA

4443 TGGCTACTTCATTCCCTACCTCCACCTGGTGGCCCATCTCCAGGACCTGGATTGGGACCC
ACTACCTGCTGCCTTCCTACTCTCAGTTGTTGCTATTTCTGACCTCGTGGGGCGTGTGGT
CTCCGGATGGCTGGGAGATGCAGTCCCAGGGCCTGTGACACGACTCCTGATGCTCTGGAC
CACCTTGACTGGGGTGTCACTAGCCCTGTTCCCTGTAGCTCAGGCTCCCACAGCCCTGGT
GGCTCTGGCTGTGGCCTACGGCTTCACATCAGGGGCTCTGGCCCCACTGGCCTTCTCTGT
[G, T]
CTGCCTGAACTAATAGGGACTAGAAGGATTTACTGTGGCCTGGGACTGTTGCAGATGATA
GAGAGCATCGGGGGGCTGCTGGGGCCTCCTCTCTCAGGTAAGTGGAATGGGGTTCCCAGG
GGGTGAGGGCTGCCATGTTGCACAAGTAGGGGAGGGTACTATTCTCATTACAGTGTATGT
GAATATTGCCCTCTGGTGTAGTACAGTACACAGCCTGCGTGGCCAACCATAGCATCCCTG
AAATGGGTCCATGGGGCAAAGAACTTGGGGCTGGGAAAGTCTGAGTGGAAGACAAAAG

5105 CCTGGCCAGACACAGACGTAGCATTCCAGTGTGCACCCTTTCCTTTGGCCTACTGGGCCC
CAAACCAGGTATCTGAGGCACCTGGTCAAAGTTCTGCTGGCTCAGGGTGCCAGAACTTTC
AGACCTTTATCTCCTCTTACCCATTAAGTGAAGCTTTAGAAAGGCCACAGTTGGTGGGCG
CCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATGGCATGAACCCGGGAGGCGG
AGCTTGCACTGAGCTGAGATCGCGCCACTGCACTTCAGCCTGGGCGACAGAGCGAGACTC
[T, C]
GTCTCAAAAAAAAAAAAAAAAAAGAAAGGCCACAGTTGCCAGAAAGAAAGGCACAAGTATGC
CTGACTCAATCTGGATCTCCAAATCCCTGCAGGCTGGTTTGGAGGTCTTTTCTGAAGGCG
GGGAGGTGGTTGAAATTAACCTTTTGGAGCCCTTTTGGGAAACCAGAGTTCTTAAGTTTAT
CCAATATTCCATGGGAGTTCCAACCTCTGAGATGATAAGTCTTCCCTCCACCCAAAA
ATGTATCTGAGCCCTCAGCCCCAGCAATAGATCACTCATGTGTATTCTTTTCTCTCTT

FIGURE 3E